

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: July 23, 2005, 22:27:07 ; Search time 1690 Seconds
(without alignments)
458.748 Million cell updates/sec

Title: US-09-973-945B-9

Perfect score: 16

Sequence: 1 gaatacatatatttc 16

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:1:
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_st:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	103	9 HS46A3F	255508 H.sapiens C
2	16	100.0	103	9 HS46A3F	255508 H.sapiens C
3	16	100.0	280	6 CQ709049	CQ709049 Sequence
4	16	100.0	280	6 CQ709049	CQ709049 Sequence
5	16	100.0	558	11 G81093	G81093 S208P6414RC
6	16	100.0	558	11 G81093	G81093 S208P6414RC
7	16	100.0	559	11 AB165691	AB165691 Bos tauru
8	16	100.0	559	11 AB165691	AB165691 Bos tauru
9	16	100.0	566	11 BV070170	BV070170 S208P6211
10	16	100.0	566	11 BV070170	BV070170 S208P6211
11	16	100.0	653	8 IMA295790	AJ295790 Isotopis
12	16	100.0	653	8 IMA295790	AJ295790 Isotopis
13	16	100.0	1022	8 AF518122	AF518122 Templeton
14	16	100.0	1022	8 AF518122	AF518122 Templeton
15	16	100.0	1413	5 BC068189	BC068189 Danio rer
16	16	100.0	1413	5 BC068189	BC068189 Danio rer
17	16	100.0	4381	9 HSM807863	BX647717 Homo sapi
18	16	100.0	4381	9 HSM807863	BX647717 Homo sapi
19	16	100.0	5728	6 AX345003	AX345003 Sequence

C	20	16	100.0	5728	6	AX345003	AX345003 Sequence
C	21	16	100.0	6059	9	AB023158	AB023158 Homo sapi
C	22	16	100.0	6059	9	AB023158	AB023158 Homo sapi
C	23	16	100.0	6059	9	AY037299	AY037299 Homo sapi
C	24	16	100.0	6059	9	AY037299	AY037299 Homo sapi
C	25	16	100.0	6992	6	CO593660	CO593660 Sequence
C	26	16	100.0	6992	6	CO593660	CO593660 Sequence
C	27	16	100.0	7733	3	DROAGPDD	D50099 Drosophila
C	28	16	100.0	7733	3	DROAGPDD	D50099 Drosophila
C	29	16	100.0	7823	6	AX278034	AX278034 Sequence
C	30	16	100.0	7823	6	AX278034	AX278034 Sequence
C	31	16	100.0	7823	6	AX323803	AX323803 Sequence
C	32	16	100.0	7823	6	AX323803	AX323803 Sequence
C	33	16	100.0	7823	6	AX344488	AX344488 Sequence
C	34	16	100.0	7823	6	AX344488	AX344488 Sequence
C	35	16	100.0	7823	6	AX346962	AX346962 Sequence
C	36	16	100.0	7823	6	AX346962	AX346962 Sequence
C	37	16	100.0	8617	3	DMU19731	U19731 Drosophila
C	38	16	100.0	8617	3	DMU19731	U19731 Drosophila
C	39	16	100.0	23501	9	AC126176	AC126176 Homo sapi
C	40	16	100.0	23501	9	AC126176	AC126176 Homo sapi
C	41	16	100.0	23579	6	AX647373	AX647373 Sequence
C	42	16	100.0	23579	6	AX647373	AX647373 Sequence
C	43	16	100.0	26246	3	CET12G3	Z68752 Caenorhabdi
C	44	16	100.0	26246	3	CET12G3	Z68752 Caenorhabdi
C	45	16	100.0	28833	9	AC096581	AC096581 Homo sapi

ALIGNMENTS

RESULT 1
HS46A3F
LOCUS
DEFINITION
H.sapiens Cpg island DNA genomic MseI fragment, clone 46a3, forward read cpg46a3.ftla.
255508
ACCESSION
255508.1 GI:1021549
VERSION
Cpg island; genomic MseI fragment.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
1 Cross, S.H., Charlton, J.A., Nan, X. and Bird, A.P.
Purification of Cpg islands using a methylated DNA binding column
Nat. Genet. 6 (3), 226-244 (1994)
JOURNAL
MEDLINE
94282070
PUBMED
8012384
REFERENCE
2 (bases 1 to 103)
Macdonald, M., Huckle, E., Wilkinson, P. and Mickle, G.
Direct Subdivision
Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgehire,
CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
Vector: pGEM-5Zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgehire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES
source
1..103
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="46a3"
/sex="male"
/tissue_type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"

ORIGIN
Query Match 100.0%; Score 16; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.1e+03;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2005, 22:32:33 ; Search time 402 Seconds
(without alignments)
252,508 Million cell updates/sec

Title: US-09-973-945B-9

Perfect score: 16

Sequence: 1 gaatcatatatttc 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7173243 seqs, 3172129809 residues

Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
20: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
21: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
22: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
23: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
24: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
25: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
26: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	16	100.0	16	9	US-09-973-945A-9
2	16	100.0	16	9	US-09-973-945A-9
3	16	100.0	16	9	US-09-973-945A-10
4	16	100.0	16	9	US-09-973-945A-10
5	16	100.0	29	9	US-09-973-945A-11
6	16	100.0	29	9	US-09-973-945A-11
7	16	100.0	29	9	US-09-973-945A-12

C	8	16	100.0	29	9	US-09-973-945A-12	Sequence 12, Appl
C	9	16	100.0	46	9	US-09-973-945A-5	Sequence 5, Appl
C	10	16	100.0	46	9	US-09-973-945A-5	Sequence 5, Appl
C	11	16	100.0	58	9	US-09-973-945A-1	Sequence 1, Appl
C	12	16	100.0	58	9	US-09-973-945A-1	Sequence 1, Appl
C	13	16	100.0	277	20	US-10-425-115-72551	Sequence 72551, A
C	14	16	100.0	277	20	US-10-425-115-72551	Sequence 72551, A
C	15	16	100.0	280	17	US-10-242-535A-53975	Sequence 53975, A
C	16	100.0	280	17	US-10-242-535A-53975	Sequence 53975, A	
C	17	16	100.0	280	18	US-10-085-783A-53975	Sequence 53975, A
C	18	16	100.0	280	18	US-10-085-783A-53975	Sequence 53975, A
C	19	16	100.0	288	19	US-10-674-124A-25880	Sequence 25880, A
C	20	16	100.0	288	19	US-10-674-124A-25880	Sequence 25880, A
C	21	16	100.0	488	13	US-10-027-632-319280	Sequence 319280, A
C	22	16	100.0	488	13	US-10-027-632-319280	Sequence 319280, A
C	23	16	100.0	488	17	US-10-027-632-319280	Sequence 319280, A
C	24	16	100.0	488	17	US-10-027-632-319280	Sequence 319280, A
C	25	16	100.0	572	13	US-10-027-632-319280	Sequence 319280, A
C	26	16	100.0	572	13	US-10-027-632-319280	Sequence 319280, A
C	27	16	100.0	572	17	US-10-027-632-319280	Sequence 319280, A
C	28	16	100.0	572	17	US-10-027-632-319280	Sequence 319280, A
C	29	16	100.0	771	13	US-10-027-632-145013	Sequence 145013, A
C	30	16	100.0	771	13	US-10-027-632-145013	Sequence 145013, A
C	31	16	100.0	771	13	US-10-027-632-145013	Sequence 145013, A
C	32	16	100.0	771	13	US-10-027-632-145013	Sequence 145013, A
C	33	16	100.0	771	17	US-10-027-632-145013	Sequence 145013, A
C	34	16	100.0	771	17	US-10-027-632-145013	Sequence 145013, A
C	35	16	100.0	771	17	US-10-027-632-145013	Sequence 145013, A
C	36	16	100.0	771	17	US-10-027-632-145013	Sequence 145013, A
C	37	16	100.0	5728	15	US-10-311-455-74	Sequence 74, Appl
C	38	16	100.0	5728	15	US-10-311-455-74	Sequence 74, Appl
C	39	16	100.0	7823	14	US-10-239-676-197	Sequence 197, App
C	40	16	100.0	7823	14	US-10-239-676-197	Sequence 197, App
C	41	16	100.0	7823	15	US-10-311-455-2033	Sequence 2033, Ap
C	42	16	100.0	7823	15	US-10-311-455-2033	Sequence 2033, Ap
C	43	16	100.0	7823	15	US-10-240-453-291	Sequence 291, App
C	44	16	100.0	7823	15	US-10-240-453-291	Sequence 291, App
C	45	16	100.0	23579	15	US-10-017-161-1909	Sequence 1909, Ap

ALIGNMENTS

US-09-973-945A-9
Sequence 9, Application US/09973945A
Patent No. US20020169297A1
GENERAL INFORMATION:
APPLICANT: University of Kentucky Research Foundation
APPLICANT: GEN, Susheng
TITLE OF INVENTION: Genetic Insulator for Preventing Influence By Another Gene
FILE REFERENCE: 050229-0297
CURRENT APPLICATION NUMBER: US/09/973, 945A
PRIOR FILING DATE: 2001-10-11
PRIOR APPLICATION NUMBER: 60/241, 735
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 16
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-973-945A-9

Query Match 100.0%; Score 16; DB 9; Length 16;
Beet Local Similarity 100.0%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAATATATATATATTC 16
1 GAATATATATATATTC 16

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2005, 22:27:43 ; Search time 1838 Seconds
(without alignments)
331.354 Million cell updates/sec

Title: US-09-973-945B-9

Perfect score: 16

Sequence: 1 GATATATATATATTC 16

Scoring table: IDENTITY_NTC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_est8:*

9: gb_est9:*

10: gb_est10:*

11: gb_est11:*

12: gb_est12:*

13: gb_est13:*

14: gb_est14:*

15: gb_est15:*

16: gb_est16:*

17: gb_est17:*

18: gb_est18:*

19: gb_est19:*

20: gb_est20:*

21: gb_est21:*

22: gb_est22:*

23: gb_est23:*

24: gb_est24:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	113	5	BQ173432 Bg_PSGR_0
2	16	100.0	113	5	BQ173432 Bg_PSGR_0
3	16	100.0	157	7	CN649902 Bg_PSGRS
4	16	100.0	157	7	CN649902 Bg_PSGRS
5	16	100.0	188	5	BU035358 CH76M16.Y
6	16	100.0	188	5	BU035358 CH76M16.Y
7	16	100.0	202	6	CF011114 OBU6C03.P
8	16	100.0	202	6	CF011114 OBU6C03.P
9	16	100.0	230	8	AO927963 RPCI-23-2
10	16	100.0	230	8	AO927963 RPCI-23-2
11	16	100.0	245	7	CN650739 Bg_PSGRS
12	16	100.0	245	7	CN650739 Bg_PSGRS
13	16	100.0	269	2	BF655072 Bg_PSGRS
14	16	100.0	269	2	BF655072 Bg_PSGRS
15	16	100.0	292	5	BK713976 BX713976
16	16	100.0	292	5	BK713976 BX713976
17	16	100.0	303	7	CN650230 Bg_PSGRS
18	16	100.0	303	7	CN650230 Bg_PSGRS
19	16	100.0	303	7	CN652060 Bg_PSGRS
20	16	100.0	303	7	CN652060 Bg_PSGRS
21	16	100.0	312	8	AO481031 RPCI-11-2
22	16	100.0	312	8	AO481031 RPCI-11-2
23	16	100.0	320	4	BI244201 Bg_PSGR_0
24	16	100.0	320	4	BI244201 Bg_PSGR_0

25	16	100.0	330	5	BP716094
26	16	100.0	330	5	BP716094
27	16	100.0	338	9	TA231B03P
28	16	100.0	338	9	TA231B03P
29	16	100.0	352	8	BZ679256 PUBGH19TD
30	16	100.0	352	8	BZ679256 PUBGH19TD
31	16	100.0	352	8	BZ66943 PUBGH19TD
32	16	100.0	352	8	BZ66943 PUBGH19TD
33	16	100.0	365	7	CN649868 Bg_PSGRS
34	16	100.0	365	7	CN649868 Bg_PSGRS
35	16	100.0	374	4	BI404261 MI-P-CP1
36	16	100.0	374	4	BI404261 MI-P-CP1
37	16	100.0	379	5	BP679193
38	16	100.0	379	5	BP679193
39	16	100.0	393	7	CN649042 Bg_PSGRS
40	16	100.0	393	7	CN649042 Bg_PSGRS
41	16	100.0	398	2	AW352596 660031E10
42	16	100.0	398	2	AW352596 660031E10
43	16	100.0	428	1	AA085241 ZN12604.T
44	16	100.0	428	1	AA085241 ZN12604.T
45	16	100.0	430	6	BY565731

ALIGNMENTS

RESULT 1
LOCUS BQ173432
DEFINITION BQ173432 113 bp mRNA linear EST 18-SRP-2002
BQ173432 Bg_PSGR_03G05.T7 Echinococcus granulosis protoscolex (full length enriched) Echinococcus granulosis cDNA clone Bg_PSGR_03G05 5', mRNA sequence.

ACCESSION BQ173432 GI:23188133

VERSION BQ173432

KEYWORDS EST

SOURCE Echinococcus granulosis

ORGANISM Echinococcus granulosis

REFERENCE Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda; Cyclophyllidae; Taeniidae; Echinococcus.

AUTHORS 1 (bases 1 to 113)

TITLE Fernandez, C. and Matzels, R.M.

JOURNAL A survey of genes expressed in Echinococcus granulosis metacestode

COMMENT Unpublished (2001)

On Apr 29, 2002 this sequence version replaced gi:20337834.

Contact: Matzels RM

Institute of Cell, Animal and Population Biology

University of Edinburgh

Ashworth Labs, West Mains Road, Edinburgh EH93JT, Scotland

Tel: +44 131 650 5511

Fax: +44 131 650 5450

Email: r.matzels@ed.ac.uk

The library was created by Dr Cecilia Fernandez in conjunction with

Prof. R. Matzels. Sequencing was also performed by Cecilia. The

sequence contained a Polya tail (trimmed)

PCR Primers

FORWARD: M13P

BACKWARD: M13R

Plate: 03 row: G column: 05

Seq primer: 77

High quality sequence stop: 113.

Location/Qualifiers

1..113

/organism="Echinococcus granulosis"

/mol_type="mRNA"

/db_xref="taxon:6210"

/clone="Bg_PSGR_03G05"

/dev_stage="larva (protoscolex)"

/clone_id="Echinococcus granulosis protoscolex (full length enriched)"

/note="Vector: pSPORT1; Site 1: SalI (5' end); Site 2: NotI (3' end); Echinococcus granulosis is a cestode parasite of dogs (definitive host) and various domestic and wild animals as well as humans (intermediate hosts). The

Query Match 100.0%; Score 16; DB 3; Length 48;